Short Communication

Loss of Antigenicity in Stored Sections of Breast Cancer Tissue Microarrays

Jennifer H. Fergenbaum,[1] Montserrat Garcia-Closas,[1] Stephen M. Hewitt,[2] Jolanta Lissowska,[3] Lori C. Sakoda,[1] and Mark E. Sherman[1]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Department of Health and Human Services, Bethesda, MD; [2]Tissue Array Research Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD; and [3]Division of Cancer Epidemiology and Prevention, Cancer Center and M. Sklodowska-Curie Institute of Oncology, Warsaw, Poland

Abstract

Immunohistochemical characterization of tumor tissues in epidemiological studies is a promising approach to identify breast cancer subtypes with distinct etiology. The recent development of the tissue microarray (TMA) technique allows for standardized, rapid, and cost-effective immunohistochemical characterization of many cases, which is critical in epidemiological studies. Sectioning paraffin blocks at different times results in loss of material, which can be reduced by preparing many sections each time a block is cut. However, data suggest that staining intensity declines in whole sections prepared from conventional paraffin blocks with storage time, resulting in false-negative results. This problem would be accentuated in TMAs because of the limited tissue representation of each case. To evaluate this concern, we prepared a single TMA block from 125 invasive breast carcinomas collected in a population-based case-control study conducted in Poland and compared estrogen receptor (ER-α), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression in sections cut and stored for 6 months at room temperature with sections cut from the same TMA block and stained on the same day. Percentage of positive cases for stored *versus* fresh sections was similar for ER (59.0%) but significantly higher in fresh sections for PR (56.3% *versus* 64.1%, $P = 0.01$) and HER2 (45.5% *versus* 64.4%, $P < 0.001$). Among cases positive in both stored and fresh sections, the median percentage of immunoreactive cells was significantly reduced and the staining intensity was consistently lower in stored compared with fresh sections. We conclude that loss of immunoreactivity is an important problem in TMAs of breast cancer. Improved methods for sectioning TMAs and storing tissue sections aimed at reducing loss of immunoreactivity are critical for the use of TMAs in epidemiological studies. (Cancer Epidemiol Biomarkers Prev 2004;13(4):667–672)

Introduction

Immunohistochemical characterization of tumor tissues in epidemiological studies is a promising approach to identify breast cancer subtypes with distinct etiology. An important limitation of previous studies has been the small size of most studies and the use of nonstandardized immunohistochemical assays performed at different study centers in larger studies (Althuis *et al.* ''Etiology of hormone receptor-defined breast cancer: a systematic review of the literature'' submitted for publication). Tissue microarrays (TMAs), which consist of recipient paraffin blocks that contain small tissue cores removed from targets of many individual donor paraffin blocks, offer an attractive method for obtaining standardized, rapid, and cost-effective immunohistochemical characterization of many tumors (1). Before histological sections of paraffin-embedded tissues can be mounted on slides for immunostaining, tissues must be repeatedly sectioned to achieve a flat surface (''facing the block''), which results in loss of specimen. Cutting and storing multiple sections every time blocks are faced conserves tissue and allows the rapid return of loaned blocks but assumes that immunoreactivity is stable over time.

Recent clinicopathological studies have reported loss of staining intensity for several markers in stored whole sections prepared from standard tissue blocks (2–5). This problem is likely to be worse in TMAs because they contain only limited tissue representation of each case. Accordingly, patchy loss of antigenicity in stored sections of TMAs may not only reduce the intensity and extent of staining but also result in unacceptably numerous false-negative results. However, repeated sectioning of the same TMA block on multiple occasions may exhaust targets, requiring exclusion of cases with inadequate material remaining in the TMA block and reducing the value of the TMA for future assays.

To assess this concern, we performed an immunohistochemical study that compared results for sections of a single TMA block from 125 invasive breast carcinomas

Received 7/29/03; revised 11/5/03; accepted 12/9/03.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Mark E. Sherman, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, 6120 Executive Boulevard, Executive Plaza South, Room 7080, MSC 7234, Rockville, MD 20852. Phone: (301) 594-7661; Fax: (301) 402-0916. E-mail: shermanm@mail.nih.gov

that were cut and stored for 6 months prior to staining with sections that were cut and stained on the same day. The cases were collected in a population-based case-control study conducted in Poland sponsored by the National Cancer Institute (Bethesda, MD).

Materials and Methods

Study Population. Breast cancer cases were participants in a population-based case-control study of breast cancer in Poland, conducted by the National Cancer Institute in collaboration with the Cancer Center and M. Sklodowska-Curie Institute of Oncology in Warsaw and the Institute of Occupational Medicine in £ódz. Appropriate institutional review boards in the United States and Poland approved the study. Eligible cases for the case-control study were women aged 19–80 years residing in Warsaw and £ódz who had been newly diagnosed with pathologically confirmed *in situ* or invasive breast cancer between February 1, 2000 and January 31, 2003. The participation rate for cases was about 80%, with tissue blocks obtained from over 90% of enrolled cases. This report analyzed a random sample of 125 invasive breast carcinoma specimens obtained from a population of patients who had not received chemotherapy, radiation, or hormonal treatment prior to tumor resection.

Specimens. The tissues included in this study were obtained from mastectomy (n = 110), quadrantectomy (n = 6), lumpectomy (n = 8), or other (n = 1) resections. Pathologists at four hospitals in Poland collected research samples of tumor that were fixed in formalin and prepared as paraffin-embedded blocks according to standard protocols. We microscopically reviewed one H&E-stained section per block to identify foci of tumor suitable for TMA construction. The tumors were classified as invasive ductal (n = 80), lobular (n = 25), mixed ductal carcinoma *in situ* and invasive ductal (n = 6), medullary (n = 2), tubular or cribriform (n = 3), papillary (n = 2), mucinous (n = 1), and other (n = 6).

TMA Construction and Preparation of Sections. We reembedded blocks in fresh paraffin and prepared TMAs using a manual arraying instrument (Beecher Instruments, Silver Spring, MD) at the Advanced Technology Center (National Cancer Institute, Gaithersburg, MD) (1). We selected targets for arraying by marking the corresponding areas on H&E-stained sections of each block. Three tissue cores (diameter 0.6 mm, height 3–4 mm) were transferred from each ''donor'' block to the ''recipient'' TMA block (45 × 20 mm, 0.7 mm center), which was designed to contain 375 cores from 125 cases in 15 subarrays (three columns of five subarrays each containing 25 cores). One of three cores per case was placed in each column to permit independent assessment. Two additional 5 × 5 subarrays of normal tissues including breast, liver, ovary, salivary gland, kidney, and breast cancer cell lines were included in the TMA block as negative and positive controls. Following TMA construction, 50 sections (5 μm) were cut using the adhesive-coated tape system and transferred to adhesive-coated slides (Instrumedics, Hackensack, NJ) and stored at room temperature for 6 months prior to immunostaining (''stored sections''). During the intervening 6 months, the TMA block was stored at room temperature. Then, sections from the TMA block were cut and immunostained on the same day (''fresh sections'') in the same staining batch as the stored sections.

Immunohistochemistry. Immunostaining for estrogen receptor (ER-α), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) was performed on the DAKO Autostainer (DAKO Corp., Carpinteria, CA) using the avidin-biotin peroxidase method (6) in one batch using a standard protocol. Sections were deparaffinized through two changes of xylene (3 min), rehydrated through graded alcohols to distilled water, and subjected to antigen retrieval using DAKO Target Retrieval Solution and steam for 20 min. After blocking endogenous peroxidase activity with hydrogen peroxidase, primary antibodies were applied: mouse monoclonal for ER (Novocastra, Newcastle Upon Tyne, UK, NCL-ER-6F11/2; 1:40 for 60 min), mouse monoclonal for PR (Novocastra NCL-PGR-312; 1:100 for 60 min), and rabbit polyclonal for HER2 (DAKO A0485; 1:125 for 30 min). Biotinylated goat anti-mouse and anti-rabbit antibody (DAKO Binotiylated Multi-Link) followed by streptavidin-horseradish peroxidase were used as secondary and tertiary reagents. Sections were visualized with diaminobenzidine (DAKO) followed by light counterstain with hematoxylin.

Pathology Evaluation for Determining Antigen Expression. Each core was evaluated for the percentage of invasive tumor cells stained on a continuous scale (0–100%) and for intensity on a three-point scale (none = 0, weak = 1, moderate = 2, strong = 3). Cores in which tumor occupied less than 10% of the total area that yielded negative immunostaining results were considered unsatisfactory and were excluded from the analysis. Stains for ER and PR were evaluated for specific nuclear staining and stains for HER2 were assessed for complete membranous staining. TMA slides were evaluated masked to time to staining (stored *versus* fresh).

Data Analysis. Results for cases having at least one satisfactory core in both stored and fresh sections were compared in paired analyses. A case was categorized as positive for a marker if staining was identified on any core; the average percentage of cells stained was calculated considering all satisfactory cores whether positive or negative (up to three cores).

Immunoreactivity as the percentage of positive cases in stored and fresh sections was compared in paired 2 × 2 tables using McNemars' test. A case with positive staining in either stored or fresh sections is considered to be a true positive case; therefore, we also calculated the percentage of positive cases for stored and fresh sections using the total number of positive cases according to either stored or fresh sections as the denominator (*i.e.*, number of stored positives/total number of positives or number of fresh positives/total number of positives). Percentage agreement was calculated as the number of pairs concordantly categorized (negative or positive) in stored and fresh sections divided by the total number of pairs multiplied by 100. Agreement between stored and fresh sections was also assessed using unweighted κ *statistics;* $\kappa \geq 0.4$ was considered a moderate and $\kappa \geq 0.8$ was considered a strong level of agreement (7). We compared the percentage of cells stained per case (based on the average of up to three cores) in stored and fresh sections using the Wilcoxon signed rank test. Staining intensity was compared for stored and fresh sections by calculating the sum of intensity scores for all satisfactory cores divided by the total number of such cores converted to a categorical variable: 0–1.0 (weak), 1.1–2.0 (moderate), and 2.1–3.0 (strong). Categorical classifications of intensity were compared as percentage agreement, unweighted statistics, and by calculating the ratio of cases showing greater staining intensity in stored as compared with fresh sections. Percentages and κ values are shown with corresponding 95% confidence intervals (CIs). The SAS Statistical Package, Version 8.1 was used for all analyses (SAS Institute, Cary, NC).

Results

Percentage of Positive Results for ER, PR, and HER2 Expression in Stored and Fresh TMA Sections. Table 1 shows the cross-classification of stored and fresh sections stained for ER, PR, and HER2. For ER, a similar percentage of cases were classified as positive in stored and fresh sections (59.0% *versus* 59.0%, P = 1.0). Significantly fewer cases were classified as PR positive in stored *versus* fresh sections (56.3% *versus* 64.1%, P = 0.01); this disparity was more pronounced for HER2 (45.5% *versus* 64.4%, P < 0.001). When the total number of positive cases by either stored or fresh sections (ER, n = 65; PR and HER2, n = 67) was considered as the denominator, the percentage of positive cases in stored sections and fresh sections was 90.8% (95% CI = 83.8–97.8) in both stored and fresh sections for ER, 86.6% (95% CI = 78.4–94.8) and 98.5% (95% CI = 95.6–100) for PR, and 68.6% (95% CI = 57.5–79.7) and 97.0% (95% CI = 93.0–100) for HER2.

Agreement between Stored and Fresh TMA Sections Cross-Classified as Negative and Positive. The percentage agreement between stored and fresh sections was higher for ER (88.0%, 95% CI = 81.6–94.4) and PR (90.3%, 95% CI = 84.6–96.0) than for HER2 (77.2%, 95% CI = 69.0–85.4; Table 1). The κ statistic indicated moderate to strong agreement for ER (κ = 0.75, 95% CI = 0.62–0.88) and PR (κ = 0.80, 95% CI = 0.68–0.92) and slightly lower agreement for HER2 (κ = 0.56, 95% CI = 0.41–0.70).

Table 1. Percentage of positive cases in stored and fresh sections stained for ER, PR, and HER2 expression in breast cancer TMAs

Stored[a]	Fresh [n (%)]		Total [n (%)]
	−	+	
ER			
−	35 (35.0)	6 (6.0)	41 (41.0)
+	6 (6.0)	53 (53.0)	59 (59.0)
Total	41 (41.0)	59 (59.0)	100 (100)
PR			
−	36 (35.0)	9 (8.7)	45 (43.7)
+	1 (1.0)	57 (55.3)	58 (56.3)
Total	37 (35.9)	66 (64.1)	103 (100)
HER2			
−	34 (33.7)	21 (20.8)	55 (54.5)
+	2 (2.0)	44 (43.6)	46 (45.5)
Total	36 (35.6)	65 (64.4)	101 (100)

Note: Any staining designated the case as positive: nuclear staining for ER and PR and complete membranous staining for HER2.
[a]Stored sections were kept for 6 months at room temperature.

Table 2. Percentage of cases having weak, moderate, and strong staining intensity among positive cases in stored and fresh sections stained for ER, PR, and HER2 expression in breast cancer TMAs

Stored[a]	Fresh [n (%)]			Total [n (%)]
	Weak	Moderate	Strong	
ER				
Weak	23 (35.4)	10 (15.4)	0 (0)	33 (50.8)
Moderate	3 (4.6)	10 (15.4)	6 (9.2)	19 (29.2)
Strong	0 (0)	1 (1.5)	12 (18.5)	13 (20.0)
Total	26 (40.0)	21 (32.3)	18 (27.7)	65 (100)
PR				
Weak	7 (10.5)	10 (14.9)	6 (9.0)	23 (34.3)
Moderate	0 (0)	2 (3.0)	11 (16.4)	13 (19.4)
Strong	0 (0)	2 (3.0)	29 (43.3)	31 (46.3)
Total	7 (10.5)	14 (20.9)	46 (68.7)	67 (100)
HER2				
Weak	44 (65.7)	8 (11.9)	0 (0)	52 (77.6)
Moderate	1 (1.5)	4 (6.0)	1 (1.5)	6 (9.0)
Strong	0 (0)	0 (0)	9 (13.4)	9 (13.4)
Total	45 (67.2)	12 (17.9)	10 (14.9)	67 (100)

[a]Stored sections kept for 6 months at room temperature.
Note: Staining intensity was assessed as the average staining intensity summed across all satisfactory cores divided by the total number of such cores converted to a categorical variable: 0–1.0 (weak), 1.1–2.0 (moderate), and 2.2–3.0 (strong).

Comparison of Diffuseness and Intensity of Staining for Stored and Fresh TMA Sections. The percentage of cells stained per case among positive cases was lower for stored sections compared with fresh sections for all markers (Fig. 1). Median percentages for stored *versus* fresh sections were 27.0% *versus* 38.0% (P = 0.001) for ER, 30.0% *versus* 70.0% (P < 0.0001) for PR, and 0% *versus* 30.0% (P < 0.0001) for HER2.

The percentage agreement for ER and PR staining intensity among positive cases was moderate (ER: 69.2%, 95% CI = 58.0–80.4; PR: 56.7%, 95% CI = 44.8–68.6) and stronger for HER2 (85.1%, 95% CI = 76.6–93.6). The κ statistic demonstrated moderate agreement for ER (κ = 0.52, 95% CI = 0.35–0.69), slight agreement for PR (κ = 0.29, 95% CI = 0.13–0.44), and stronger agreement for HER2 (κ = 0.66, 95% CI = 0.48–0.85). The ratio of cases showing stronger intensity in stored as compared with fresh sections was <1.0 for all markers (ER = 4/16, PR = 2/27, and HER2 = 1/9; Table 2; Fig. 2).

Discussion

Our data indicate that sections of breast cancer TMAs stored for 6 months at room temperature show reduced immunostaining for ER, PR, and HER2 compared with freshly cut sections stained on the same day taken from the same TMA block. The greatest loss in the percentage of immunoreactive cases in stored sections was found for HER2; the percentage of cells stained and staining intensity were dramatically reduced for all markers in stored

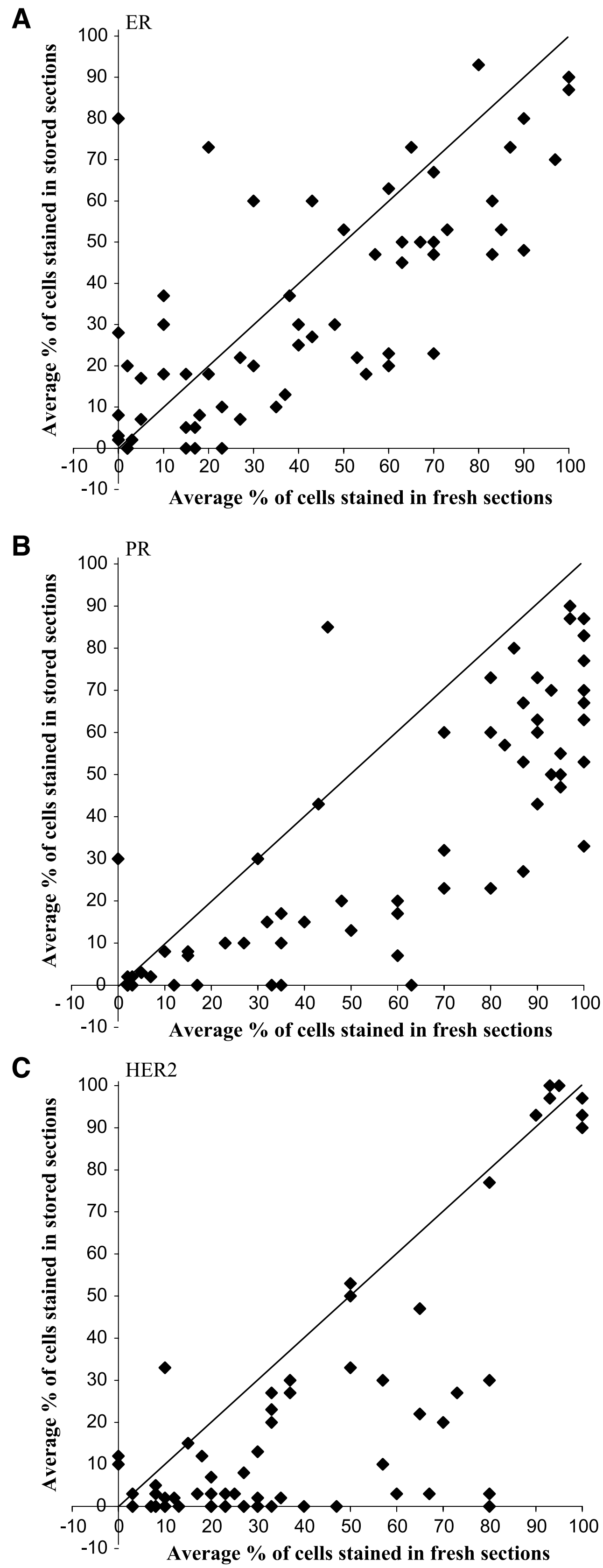

Fig. 1. Average percentage of cells stained per case in stored *vs.* fresh sections among positives for **A.** ER (n=65), **B.** PR (n=67), and **C.** HER2 (n=67).

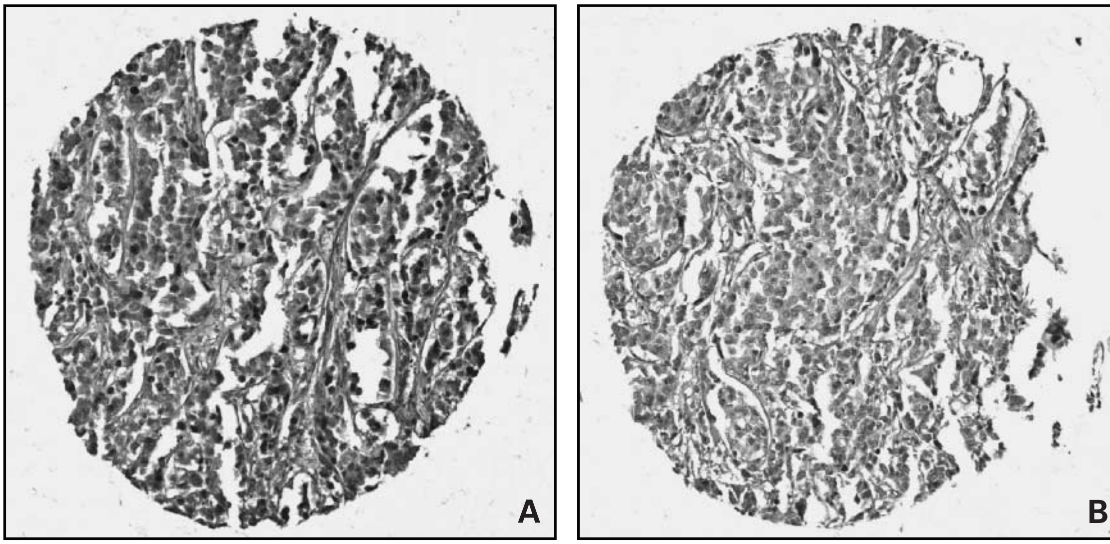

Fig. 2. Immunohistochemistry for PR. **A.** A core from the freshly cut and immunostained section. **B.** The same tissue core from a section that had been cut 6 months previously and was stained at the same time as the core in **A**.

sections, with the most striking losses found for PR. Overall, extent and intensity of staining were reduced more by storage than classification of cases as negative or positive.

Previous investigations have observed a loss of staining intensity for ER following storage of sections (2–4) and have demonstrated that increased length and higher storage temperatures increase loss of antigenicity for PR and HER2. However, cold storage does not eliminate antigen degradation (5). Loss of immunoreactivity has been demonstrated with both light microscopic and computer-generated interpretation of stains. One report has suggested that nuclear and cytoplasmic stains are less affected by storage time and temperature than membranous stains (5). The cause of antigen degradation over time is unknown, but oxidation has been suggested as a mechanism and storage under nitrogen may preserve antigenicity.

For many markers, the etiological and clinical relevance of minimal staining is not entirely certain (8, 9) and cut points for scoring stains as positive have not been standardized. In this study, we evaluated staining of any extent and intensity because we could not exclude that immunoreactivity below cut points set for determining therapeutic options (*i.e.*, Herceptin, tamoxifen, etc.) might prove to have etiological relevance. For example, had we set moderate intensity as a threshold for a positive HER2 stain, 29.6% of cases in fresh sections would have been classified as positive, similar to that reported for 2+ or 3+ staining using standardized clinical protocols.

Although we stained stored and fresh sections simultaneously using the standard protocol used in our institution, we recognize that results of immunohistochemical stains obtained with protocols used in other laboratories may differ. We also recognize that the paraffin that had been used in Poland differed from that used in the United States and that routine deparaffinization may not completely remove paraffin from tissue sections. However, we performed pilot work that demonstrated that residual paraffin was unrelated to our staining results (data not shown).

In our study, 17.6–20.0% of cases lacked a satisfactory core when the TMA was faced and sectioned a second time (fresh sections, cut after 50 sections) compared with 3.2–4.8% of cases at the initial sectioning (stored sections). This highlights the loss of tissue when a block is refaced and cut again. Similarly, losses of 10% to over 30% have been reported for different tissue types despite the routine inclusion of multiple cores per case (10). These data underscore the need to consider tissue access, tissue heterogeneity, redundancy of tissue core representation, and preparation of multiple TMA blocks in planning epidemiological investigations.

Efforts to develop improved microtome sectioning techniques that maximize tissue conservation and validation of improved methods for preserving immunoreactivity in stored tissue sections are urgently needed. Currently, we are working with storage under nitrogen and shipping in vacuum-packed bags. As high throughput quantitative techniques for assessing immunostaining assume greater importance in etiological and translational research, objective quality assurance guidelines will likely become a growing concern.

Acknowledgments

We thank the pathology department at the following institutions for the collection of tissue samples included in this study: Cancer Center and M. Sklodowska-Curie Institute of Oncology (Warsaw, Poland); Oncological Clinic of Medical Academy (£ódz, Poland); and Polish Mother's Memorial Hospital (£ódz, Poland). We also thank Kimberly Parker at Advanced Technology Center (Gaithersburg, MD) for technical help. Supported in part by a National Cancer Institute Breast Cancer Faculty Award (M.E.S.).

References

1. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med, 1998;4:844–7.

2. Wester K, Wahlund E, Sundstrom C, et al. Paraffin section storage and immunohistochemistry. Effects of time, temperature, fixation, and retrieval protocol with emphasis on p53 protein and MIB1 antigen. Appl Immunohistochem Mol Morphol, 2000;8:61–70.
3. Bertheau P, Cazals-Hatem D, Meignin V, et al. Variability of immunohistochemical reactivity on stored paraffin slides. J Clin Pathol, 1998;51:370–4.
4. Jacobs TW, Prioleau JE, Stillman IE, Schnitt SJ. Loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. J Natl Cancer Inst, 1996;88:1054–9.
5. Van den Broek LJ, Van de Vijver MJ. Assessment of problems in diagnostic and research immunohistochemistry associated with epitope instability in stored paraffin sections. Appl Immunohistochem Mol Morphol, 2000;8:316–21.
6. Hsu SM, Raine L, Fanger H. Use of avidin-biotin-peroxidase complex (ABC) in immunoperoxidase techniques: a comparison between ABC and unlabeled antibody (PAP) procedures. J Histochem Cytochem, 1981;29:577–80.
7. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
8. Layfield LJ, Gupta D, Mooney EE. Assessment of tissue estrogen and progesterone receptor levels: a survey of current practice, techniques, and quantitation methods. Breast J, 2000;6:189–96.
9. Hanna W. Testing for HER2 status. Oncology, 2001;61 Suppl 2: 22–30.
10. Hoos A, Cordon-Cardo C. Tissue microarray profiling of cancer specimens and cell lines: opportunities and limitations. Lab Invest, 2001;81:1331–8.